

# Factor analysis of chemical composition in the Karoon River basin, southwest of Iran

Heidar Zarei · Mohsen Pourreza Bilondi

Received: 12 March 2013 / Accepted: 14 August 2013 / Published online: 8 September 2013  
© The Author(s) 2013. This article is published with open access at Springerlink.com

**Abstract** There are a number of different parameters like discharge which affect water quality and conditions. Typically, water quality is expected to improve by the increase in discharge and significantly influenced by the runoff water quality entering the river system. As a result this research subjected to address the increase of salinity in the wet season. To reach this point, a multivariate statistical technique, namely factor analysis (FA), has been used to assess the spatial variability of water quality parameters and conditions in the Karoon River basin located in southwest Iran. The various water quality parameters (10 variables) and discharge were incorporated to FA to better interpret the processes (natural) and specific source of water quality deterioration. The results revealed that water quality variations are affected mostly by dissolved mineral salts along the entire Karoon River. Furthermore, major contamination threat is caused by geological situation for over the year which is defined as nonpoint pollution source and may explain most part of the observed variances (50 %) in the data. Then as a general result of this study it can be claimed that spatially and temporally management of water use in different parts can be carried out effectively by FA technique.

**Keywords** Water quality · Factor analysis · Water chemical composition · Karoon River · Geological situation

## Introduction

Many water quality variables are subject to large fluctuations in space and time. Understanding these fluctuations in the physical environment and determining whether such changes are natural or a result of anthropogenic influences can be a challenging task. An ideal variable provides unambiguous information about the condition of the environment in relation to reference conditions and is relatively easy and inexpensive to measure.

The chemical composition of the surface and groundwater is controlled by many factors that include composition of precipitation, mineralogy of the watershed and aquifers, climate, and topography. These factors combine to create diverse water types that change spatially and temporally (Guler et al. 2002). The usual program of water quality assessment is measuring of multiple parameters in different monitoring sites in specific period. Therefore, a complex data matrix should be assessed to evaluate water quality (Chapman 1992). Furthermore, in river monitoring, it is frequent to determine whether a variable should be attributed to man-made (spatial, anthropogenic) pollution or natural (temporal, climatic) changes in the river hydrology (Razmkhah et al. 2010). Several authors already have reported the use of water quality indices (WQI) to verify the influence of waste discharges on water quality of rivers. Despite the fact that WQI provides a useful way to predict changes in water quality, considering multiple parameters, it does not provide evidences on the contamination sources, especially because they are calculated after

H. Zarei  
Department of Hydrology and Water Resources, Shahid Chamran University of Ahvaz, Ahvaz, Iran  
e-mail: zareih@scu.ac.ir

M. Pourreza Bilondi (✉)  
Department of Water Engineering, University of Birjand, Birjand, Iran  
e-mail: Mohsen.pourreza@birjand.ac.ir

normalization of analytical values (Pesce and Wunderlin 2000). The usual technique of interpretation of river water quality is a univariate procedure which does not adequately characterize simultaneous similarities and differences between samples or variables (Dixon and Chiswell 1996). The problem of environmental data reduction and interpretation of multiconstituent physicochemical and biochemical measurements can be more easily handled through application of robust multivariate methods and exploratory data analysis (Brown 1988; Wenning and Erickson 1994). Factor analysis (FA) as a data reduction technique is widely used, being capable of detecting similarities among samples and/or variables (Wenning and Erickson 1994; Battegazzore and Renoldi 1995; Voutsas et al. 1995; Wang et al. 2006; Mendiguchía et al. 2007). In summary, the goal is to explain a portion of their variance in the set of variables input into the analysis by identifying certain underlying common dimensions called the factors.

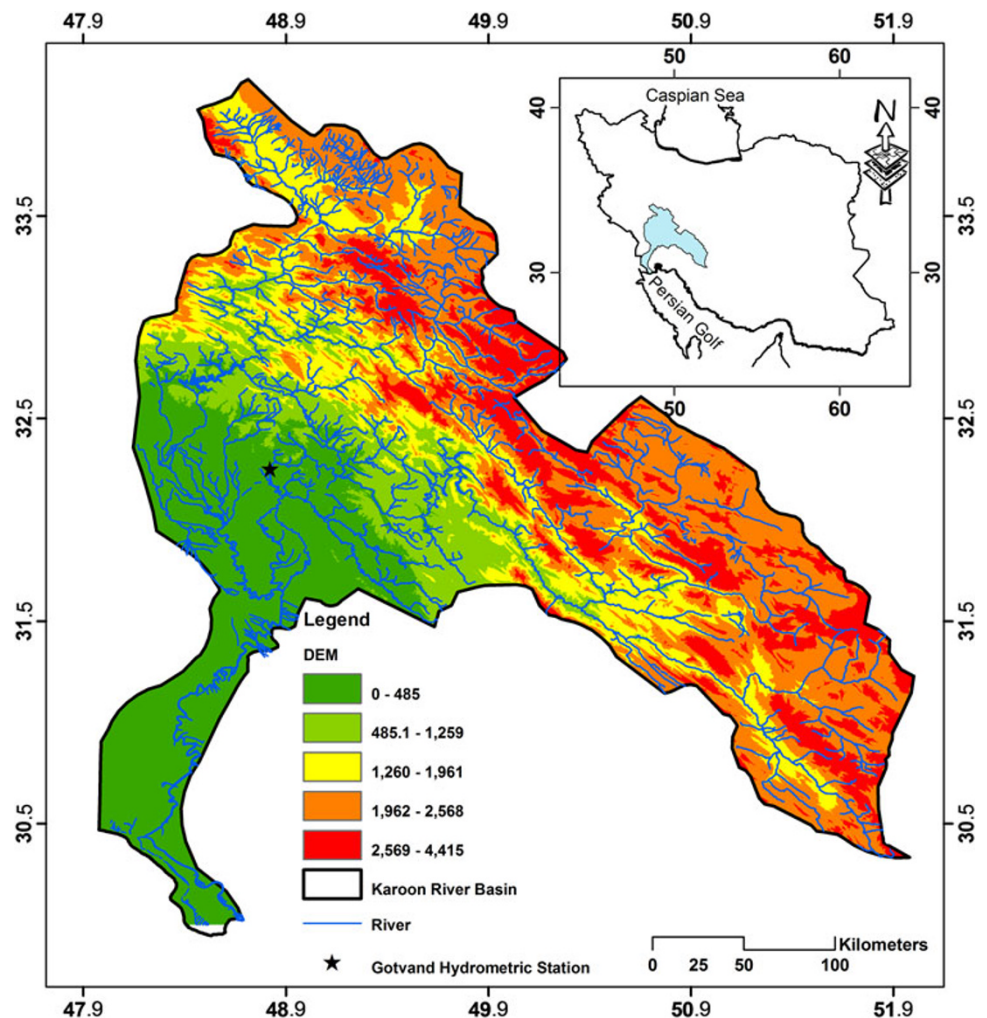
The main objective of this study is to identify the source of surface contamination in the Karoon river basin using FA as an effective multivariate statistical technique. The

water quality data were incorporated to FA to better interpret, understand and define the mechanisms, processes (natural) and specific source of water quality deterioration and contamination in the basin. Selection of this river was done for some reasons; The Karoon River is the largest river in Iran which irrigates more than 280,000 hectares of the surrounding plain. Consequently better understanding of the river water quality and its dominate controlling factors are crucial for providing the basis for comprehending river basin processes and it is essential for developing future projects in agricultural, urbanization and industrial sectors as well as water resource planning and management.

### Study area

The Karoon River basin is located in southwestern Iran; it is the largest river in Iran with an area of 42,574 km<sup>2</sup> (Fig. 1) The Karoon river system supplies the water demands of sixteen cities, several villages, thousands of

**Fig. 1** Sketch map of water quality monitoring station of Karoon River basin in Iran



hectares of agricultural lands, and several hydropower plants. Increasing water demands at the development stage including agricultural networks, fish hatchery projects, and inter-basin water transfers could result in a gloomy future for water quality of the Karoon River (Karamouz et al. 2004).

The most widespread of the outcrops in the basin is related to the limestone formations (more than 70 % of basin area). Another important formation (<10% of basin area) in the basin is named “Gachsaran”. The lithological composition of Gachsaran Formation mainly consists of marl, anhydrite, gypsum, halite and some interbeds of marly limestone.

## Materials and methods

### Monitored parameters

The data were obtained from the hydrological database of Khuzestan Water and Power Authority, Ahvaz, Iran over the period of 1976–2005. Khuzestan Water and Power Authority is the most important organization concerned about water resources in Iran since most of surface water resources are located in this province. The quality of data is usually very good because of the high quality of instruments they used in the stations. Besides, there is a team-work community to control the quality of collected data in this organization. The discharge and the water quality parameters have been monitored at the Gatvand station on a monthly basis for a period of 30 years. This station was selected for the purpose of this study, since it is located before the Khuzestan plain, where the Karoon River enters the plain and thus is less affected by agricultural activities in the region. The selected water quality parameters include calcium, magnesium, bicarbonates, sulfate, chlorides, electrical conductivity, total dissolved solid, pH and discharge. These parameters were chosen since they have long-term records of measured data. Based on the results of the previous studies the flow discharge logarithm was used instead of discharge measurements to obtain more effective results (Evansa et al. 1996).

Due to the fact that the hydrometric station has not been regularly measured and sampled during all the months of a year and also to remove the potential impacts of the seasonal changes, the input data were transformed into seasonal data. The water quality parameters, their units and the basic statistics of the data set on river water quality for the four seasons are summarized in Table 1. A well-defined temporal heterogeneity in distribution of different water quality parameters was observed in the studied station. For example, the box plot diagram of electrical conductivity is shown in Fig. 2. The box plot diagram identified was

constructed to evaluate different patterns associated with temporal variations in river water quality. Based on the box plot diagram of electrical conductivity, certain water quality constituents differ seasonally.

The monitoring data for various water quality parameters (10 variables) and discharge were organized in a matrix with variables as columns and the samples (289 cases) as rows. SPSS software was used to analysis the statistical relationships of the data after performing auto scaling for various parameters. To examine the suitability of the data for FA, Kaiser–Meyer–Olkin (KMO) and Bartlett’s test were performed. KMO is a measure of sampling adequacy that indicates the proportion of variance which is common variance, i.e., which might be caused by underlying factors (Shrestha et al. 2008). High values (close to 1) typically indicate that FA may be useful, which are the case in this study shown in Table 2.

Bartlett’s test of sphericity indicates whether the correlation matrix is an identity matrix, which would indicate that variables are not specified in confidence interval. The significance level, which is 0.00 in this study (<0.05), indicates that there are significant relationships among variables. The water quality parameters were grouped into four seasons: autumn (October–December) and winter (January–March), spring (April–June) and summer (July–September).

### Factor analysis (FA)

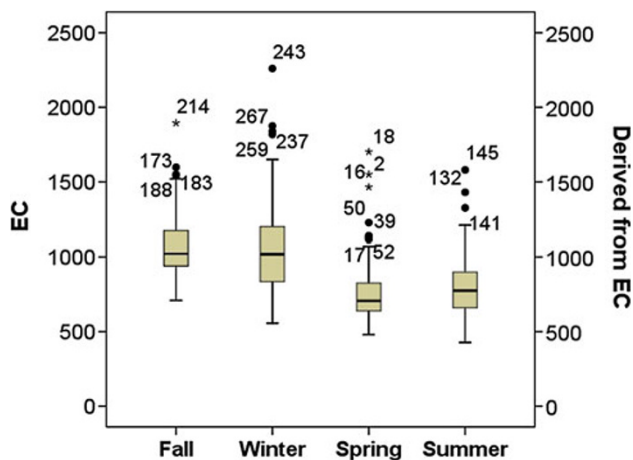
Multivariate statistical techniques, such as FA or principal component analysis (PCA), provide more insight into the underlying structure of a data set, the use of these techniques might require further analyses to identify distinct groups (Belkhiri and Mouni 2012). FA is related to PCA, but the two are not identical. FA as a data reduction technique is performed in two steps included of extraction and rotation of factors. The main purpose of first step is to extract various factors through PCA. Application of PCA has been verified by many scientists to water quality assessment (Gotz et al. 1998; Da Silva and Sacomani 2001; Simeonov et al. 2003; Parinet et al. 2004; Ouyang 2005; Kuppusamy and Giridhar 2006; Nakano et al. 2008). For example, Belkhiri et al. (2011) adopted R-mode FA and their results indicate that the parameters responsible for groundwater quality variations are mainly related to the presence and dissolution of some carbonate, silicate, and evaporite minerals in the aquifer.

In second step, rotation of factors can achieve a simpler and more meaningful representation of the underlying factors by decreasing contributions to principal components by variables with minor significance and increasing the more significant ones (Razmkhah et al. 2010).

**Table 1** Descriptive statistics of water quality parameters in Gotvand station for different seasons

Season	Variable unit	Discharge (Q) m <sup>3</sup> /s	Total dissolved solid (TDS) mg/l	Electrical conductivity (EC) $\mu$ S/cm	Level of acidity (pH)	Bicarbonate (HCO <sub>3</sub> ) meq/l	Chloride (Cl) meq/l	Sulfate (SO <sub>4</sub> ) meq/l	Calcium (Ca) meq/l	Magnesium (Mg) meq/l	Sodium (Na) meq/l	Potassium (K) meq/l
Autumn	N											
	Valid	68	69	69	69	69	69	69	69	69	69	69
	Missing	1	0	0	0	0	0	0	0	0	0	0
	Mean	316	674	1,079	8.1	2.67	5.82	2.13	3.45	1.53	5.64	0.05
	Median	246	634	1,022	8.1	2.70	5.50	1.90	3.15	1.40	5.30	0.04
	Mode	264	610 <sup>a</sup>	1,020	8.3	2.65 <sup>a</sup>	5.40	1.00 <sup>a</sup>	2.70	1.35	5.10 <sup>a</sup>	0.04
	SD	285	157	228	0.3	0.34	1.67	1.18	1.06	0.56	1.66	0.02
	Cv	0.90	0.23	0.21	0.04	0.13	0.29	0.55	0.31	0.37	0.29	0.40
	Minimum	62	453	710	7.3	1.67	2.90	0.61	2.05	0.44	2.60	0.02
	Maximum	2,217	1,238	1,895	8.7	3.40	11.57	7.40	8.25	3.20	11.50	0.10
Winter	N											
	Valid	72	74	74	74	74	74	74	74	74	74	74
	Missing	2	0	0	0	0	0	0	0	0	0	0
	Mean	666	676	1,070	8.1	2.60	5.48	2.55	3.82	1.41	5.42	0.05
	Median	419	645	1,019	8.1	2.64	5.13	2.20	3.63	1.30	5.13	0.05
	Mode	427	410 <sup>a</sup>	900 <sup>a</sup>	8.0	2.90	2.80	2.60	3.40 <sup>a</sup>	1.20	5.30	0.05
	SD	601	205	326	0.3	0.59	2.43	1.37	1.15	0.49	2.53	0.02
	Cv	0.90	0.30	0.30	0.04	0.23	0.44	0.54	0.30	0.35	0.47	0.40
	Minimum	2	356	557	7.6	0.66	2.18	0.25	1.70	0.60	2.05	0.01
	Maximum	3,020	1,338	2,260	9.6	4.00	14.05	8.50	9.90	3.20	14.64	0.20
Spring	N											
	Valid	70	71	71	71	71	71	71	71	71	71	71
	Missing	1	0	0	0	0	0	0	0	0	0	0
	Mean	754	479	765	8.0	2.46	3.42	1.58	2.90	1.18	3.40	0.03
	Median	556	444	706	8.0	2.50	2.90	1.40	2.88	1.12	2.80	0.03
	Mode	455 <sup>a</sup>	340 <sup>a</sup>	662 <sup>a</sup>	7.9	2.00 <sup>a</sup>	1.85 <sup>a</sup>	1.05 <sup>a</sup>	3.00	1.00	2.80	0.03
	SD	654	151	235	0.2	0.44	1.79	0.83	0.71	0.29	1.88	0.01
	Cv	0.87	0.32	0.31	0.03	0.18	0.52	0.53	0.24	0.25	0.55	0.33
	Minimum	72	240	480	7.6	1.39	1.40	0.60	1.80	0.51	1.53	0.02
	Maximum	4,487	1,091	1,702	8.6	3.40	11.50	5.48	5.80	2.20	13.00	0.06
Summer	N											
	Valid	75	75	75	75	75	75	75	75	75	75	75
	Missing	0	0	0	0	0	0	0	0	0	0	0
	Mean	352	491	790	8.1	2.47	3.85	1.28	2.64	1.19	3.71	0.03
	Median	307	480	775	8.1	2.50	3.70	1.25	2.70	1.14	3.60	0.03
	Mode	160 <sup>a</sup>	530	800	8.1	2.50 <sup>a</sup>	2.15 <sup>a</sup>	0.95 <sup>a</sup>	3.00	0.90 <sup>a</sup>	3.20 <sup>a</sup>	0.03
	SD	205	128	202	0.2	0.38	1.56	0.56	0.44	0.34	1.49	0.01
	Cv	0.58	0.26	0.26	0.02	0.15	0.41	0.44	0.17	0.29	0.40	0.33
	Minimum	106	261	428	7.6	1.60	0.65	0.10	1.40	0.62	0.65	0.02
	Maximum	1,000	1,030	1,580	8.7	3.40	10.95	4.20	3.65	2.70	10.65	0.09

<sup>a</sup> Multiple modes exist. The smallest value is shown



**Fig. 2** Seasonally box plot of water quality parameters in Gotvand station

**Table 2** Kaiser-Meyer-Olkin and Bartlett's test

	Autumn	Winter	Spring	Summer
KMO	0.52	0.57	0.68	0.51
Bartlett's Test	0.00	0.00	0.00	0.00

The basic purpose of a FA are: to determine how many factors are required to provide a set of variables, to specify the level of correlation between variables and associated statistical factors, to investigate and analyze the common factors and to quantify each factor possessed by each observation (identified by the factor scores).

In fact, the main objective of the FA is to analyze the existing structure in matrix of variance–covariance of the multivariate data. The Eigenvalues and Eigenvectors are also extracted from the correlation and covariance matrix. This analysis is a multivariate statistical method which is mainly used for analyzing the relation between a set of variables or issues (Reghunath et al. 2002). This method assumes that the relation between a number of variables is reflected on a lower level called factor (Davis 1973). Thus, if the main data are replaced by the factors, the complexity of the data contents is reduced and then the factors can be assigned to the real effective chemical processes of the water flow of the basin (Evansa et al. 1996). It should be noted that for a better interpretation the initial factors should be modified by varimax rotation (Grande et al. 1996). As a result, a small number of factors will usually account for approximately the same amount of information as does the much larger set of original observations. The FA can be expressed as:

$$z_{ji} = a_{f1}f_{1i} + a_{f2}f_{2i} + a_{f3}f_{3i} + \dots + a_{fm}f_{mi} + e_{fi} \quad (1)$$

where  $z$  is the component score,  $a$  is the component loading,  $f$  is the factor score,  $e$  is the residual term accounting for errors or other source of variation,  $i$  is the sample number and  $m$  is the total number of variables.

**Table 3** Variables and factor loadings after varimax rotation for autumn season

Variable	Factor 1	Factor 2	Factor 3
Log Q		0.91	
TDS	0.90		
EC	0.93		
Ph			−0.70
HCO <sub>3</sub>			0.83
Cl	0.98		
SO <sub>4</sub>	0.41	0.84	
Ca	0.43	0.75	
Mg			
Na	0.97		
K	0.54		
Eigenvalue	4.28	2.59	1.34
% of total variance	38.87	23.58	12.21
Cumulative %	38.87	62.45	74.66

Extraction method: principal component analysis. Rotation method: Varimax with Kaiser Normalization

The FA was less used for temporal analysis of hydro-chemical studies; however, it is more applied for spatial analysis of the hydrochemical data such as identification of aeration sources using different resolved materials in different places (Reeder et al. 1972; Puckett and Bricker 1992) and determination of ions sources across a dump area (Hooper and Peters 1989).

FA was performed on the normalized data sets separately, to compare the compositional pattern between analyzed water samples and identify the factors influencing each one. The input data matrices (variables-cases) for FA were [11–289] for the Gotvand Hydrometric station.

## Results and discussion

An examination of the Kaiser-Meyer-Olkin measure of sampling adequacy suggested that the sample was factorable (Table 2). PCA was carried out to extract the various factors. The varimax rotation was performed to secure increased principal components of chemical/environmental significance. The Eigenvalues for different factors, percentage variance accounted, cumulative percentage variance and component loadings (Varimax rotated) are given in Table 2. Eigenvalues >1 were taken as criterion for the extraction of the principal components required for explaining the sources of variances in the data. This analysis resulted in the explanation of 71.53–79.26 of variances in the data.

The parameter loading for the three components from the PCA of the data set are given in Tables 3, 4, 5 and 6. In



**Table 4** Variables and factor loadings after varimax rotation for winter season

Variable	Factor 1	Factor 2	Factor 3
Log Q	−0.57	0.44	
TDS	0.92		
EC	0.94		
pH		−0.61	
HCO <sub>3</sub>			−0.90
Cl	0.96		
SO <sub>4</sub>	0.48	0.71	
Ca		0.90	
Mg	0.64		
Na	0.97		
K			0.73
Eigenvalue	4.68	2.14	1.62
% of total variance	42.54	19.47	14.72
Cumulative %	42.54	62.02	76.74

Extraction method: principal component analysis. Rotation method: Varimax with Kaiser Normalization

**Table 5** Variables and factor loadings after varimax rotation for spring season

Variable	Factor 1	Factor 2	Factor 3
Log Q		−0.90	
TDS	0.70	0.65	
EC	0.69	0.70	
pH			0.88
HCO <sub>3</sub>			−0.80
Cl	0.54	0.80	
SO <sub>4</sub>	0.92		
Ca	0.76		−0.51
Mg	0.57		
Na	0.55	0.78	
K	0.60		
Eigenvalue	3.70	3.09	1.93
% of total variance	33.61	28.08	17.58
Cumulative %	33.61	61.68	79.26

Extraction method: principal component analysis. Rotation method: Varimax with Kaiser Normalization

order to interpret the results, the high factor load in excess of 0.75 and mean factor load between 0.4 and 0.75 were considered which were optional values (Childs 1970). These values were obtained in studies by different researchers (Miller and Drever 1977; Puckett and Bricker 1992; Evansa et al. 1996).

The investigation of the factors with Eigenvalues >1 for the four seasons demonstrated that three factors affected the chemical composition of the Karoon River water quality. The Scree plot of spring season exhibited in Fig. 3,

**Table 6** Variables and factor loadings after varimax rotation for summer season

Variable	Factor 1	Factor 2	Factor 3
Log Q	−0.69		
TDS	0.89		
EC	0.88		
pH			−0.61
HCO <sub>3</sub>			0.81
Cl	0.91		
SO <sub>4</sub>		0.91	
Ca			0.77
Mg		0.58	
Na	0.89		
K			
Eigenvalue	3.98	2.12	1.77
% of total variance	36.20	19.27	16.06
Cumulative %	36.20	55.47	71.53

Extraction method: principal component analysis. Rotation method: Varimax with Kaiser Normalization

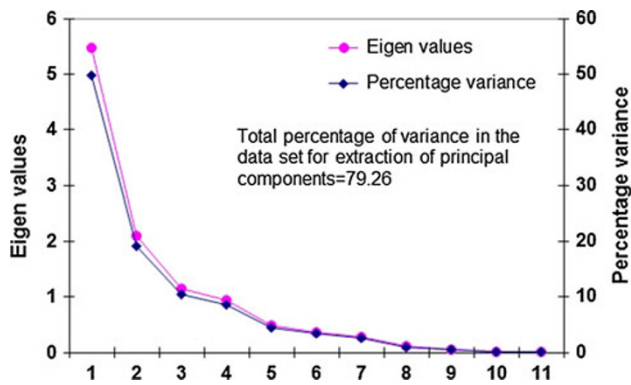
which also includes the percentage variances explained by each component and gives an idea on how the different principal components were extracted. This figure shows a pronounced change of slope after the 3th Eigenvalue. Therefore, three components were retained, which have Eigenvalues >1 and explain 79.26 % of the variance.

An Eigenvalue gives a measure of the significance of the factor: the factors with the highest Eigenvalues are the most significant. Eigenvalues of 1.0 or greater are considered significant. Liu et al. (2003) classified the factor loadings as ‘strong’, ‘moderate’ and ‘weak’, corresponding to the absolute loading values of >0.75, 0.75e0.50 and 0.50e0.30, respectively. Therefore, we also classified these results in Tables 3, 4, 5 and 6.

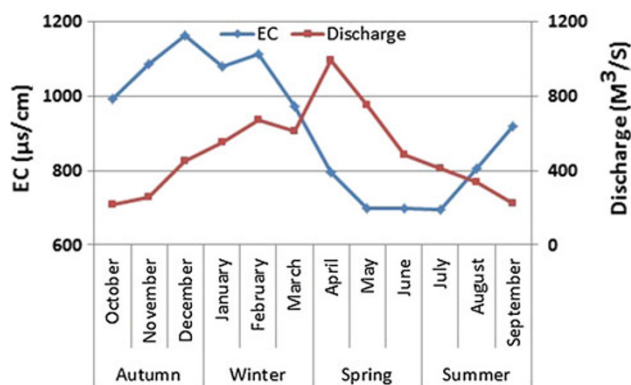
All three factors are examined separately as follows:

Autumn season: The results of the analysis discovered that three factors accounted for 74.66 % of the total variance (Table 3). Based on the distribution of the Eigenvalues, factor 1 alone explained 38.87 % of the variance. TDS, conductivity, chloride and sodium were strongly correlated and sulfate, calcium and potassium were moderately correlated with factor 1. Calcium, sulfate, and discharge were strongly correlated with factor 2 and bicarbonate and pH with factor 3.

Winter season: The results of the analysis revealed that three factors account for 76.74 % of the total variance (Table 4). The Eigenvalues also performed that factor 1 alone accounts for 42.54 % of the variance. TDS, conductivity, chloride and sodium were strongly and sulfate, magnesium and discharge were moderately correlated with factor 1. Calcium was strongly correlated; sulfate, pH and



**Fig. 3** Scree plot of Eigenvalues vs. components along with percentage variances vs. components for Gotvand water quality station



**Fig. 4** Discharge and electrical conductivity variations during a year

discharge were moderately correlated with factor 2; and bicarbonate and potassium with factor 3.

**Spring season:** The results of the analysis revealed that three factors account for 79.26 % of the total variance (Table 5). The Eigenvalues also showed that factor 1 alone accounts for 33.61 % of the variance. Sulfate, calcium and discharge were strongly and TDS, conductivity, chloride, sodium, magnesium, potassium were moderately correlated with factor 1. Discharge, chloride, sodium were strongly correlated, TDS and conductivity were moderately correlated with factor 2, and bicarbonate and pH with factor 3.

**Summer season:** The results of the analysis exhibited that three factors account for 71.53 % of the total variance (Table 6). The Eigenvalues also showed that factor 1 alone accounts for 36.20 % of the variance. TDS, conductivity, chloride and sodium were strongly and discharge was moderately correlated with factor 1. Sulfate was strongly correlated and magnesium was moderately correlated with factor 2 and calcium, bicarbonate and pH with factor 3.

Factor 1 have high loading of the ions  $\text{Na}^+$  and  $\text{Cl}^-$ , and factor 2 have great loading of the ions  $\text{Ca}^{2+}$  and  $\text{SO}_4^{2-}$  for all seasons except spring. The concentration of Na, Cl, Ca and  $\text{SO}_4$  in runoff generated by Gachsaran Formation outcrops is much greater than that in other formations. Factor 3 shows the influence of limestone formations on the Karoon River water quality.

As identified by a plot, the selected parameters showing seasonal trends are given in Fig. 4. The average discharge (Fig. 4) is higher in spring compared to autumn, winter and summer. In the study period, these might have been due to the frequent snow melt and continuously discharge melt water into river network in the basin. Lastly, river water quality is expected to improve in the Karoon River basin by increasing discharge amount. Vary factors obtained from FA indicate that the parameters responsible for water quality variations are mainly related to discharge amount as well as Gachsaran geology and formation. EC is extremely less in the spring season when discharge value increases and tremendously high in the rest of the year. This pollution is significantly natural and point source as nonpoint pollutions like agriculture and orchard plantations did not involve in this basin.

Figure 4 showed that although the lowest flow discharge was observed during the summer season, the highest salinity occurred over the autumn season. In other words, the salinity of the river water increases in the wet season (autumn and winter).

Gachsaran Formation having a high erosion potential, so the generated runoff in wet season passing through this formation, dissolve gypsum and halite, can contaminate the surface or subsurface water resources and the salinity of the river water increases. As regards in the flowing water, the dissolution rates of gypsum is 100 times more than the rate of limestone dissolution, and only about 1/1,000 the rate of salt dissolution (Milanovic 2004), so dissolution of halite of Gachsaran Formation is the first and dissolution of gypsum is the second effective factors on the Karoon River water quality.

## Conclusions and future work

In this case study, FA was used to evaluate temporal variations in water quality of the Karoon River using 30 years datasets of 10 water quality variables and discharge in the Gotvand station which is located before the Khuzestan plain and less affected by agricultural activities. Our analysis grouped 1 sampling site into four seasons for understanding temporal variation in water quality characteristics. Based on obtained information, it is possible to design a future, optimal sampling strategy, which could increase the number of sampling stations and associated

costs particularly in the agricultural and industrial areas. Although the FA/PCA would not result in a significant data increasing, it will help to identify the factors/sources responsible for variations in river water quality at different stations with different pollutions sources as well as sampling size.

Results obtained from FA indicate that the parameters responsible for water quality variations are mainly related to discharge and geology formation (natural) and not any man-made pollution sources. Future research should focus and address man-made pollution sources and the areas with highly polluted areas in the basin.

Due to the fact that in this basin only the Gachsaran Formation has saline units, it can be concluded that the most influential factor on the chemical composition of the water quality is the solution of saline units in Gachsaran Formation. The second affecting factor on the water chemical composition of Karoon River in all seasons except spring is the high-to-mean factor loads Ca and SO<sub>4</sub>. The ions are increased in result of solution of gypsum formations; however, due to their less solubility, they are the second affecting factor. Thus, major contamination threats over the year are geological situation which are defined as nonpoint pollution source. Due to less solubility of the formations in the basin compared to Gachsaran Formation, the most impact on water quality is related to Gachsaran Formation. Although Gachsaran Formation is outcropped solely in a minor part of the basin, because of its specific lithology characteristics, this formation can affect Karoon River water quality significantly (Zarei et al. 2005). This formation also causes a high erosion potential, so the generated runoff passing through this formation, dissolve gypsum and halite, can contaminate the surface or subsurface water resources. Water quality can be affected by the formation in two ways; firstly runoff water is affected by this lithology and then brine springs which flow through this formation. This result may be affected partly by studying different period as well as land use change, urbanization, etc.

Since water quality assessment is an important part of watershed river management, this study aimed to illustrate the usefulness of a statistical techniques named factor analysis. This technique help efficiently for analysis and interpretation of special data sets, and in water quality assessment, identification of pollution sources/factors and understanding seasonal variations in water quality for effective water quality river management.

**Acknowledgments** Khozestan Water & Power Authority (KWPA) is deeply acknowledged for providing the database. Also the authors sincerely wish to thank anonymous reviewers to review the manuscript.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use,

distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Battegazzore M, Renoldi M (1995) Integrated chemical and biological evaluation of the quality of the River Lambro (Italy). *Water Air Soil Pollut* 83:375–390
- Belkhiri L, Mouni L (2012) Hydrochemical analysis and evaluation of groundwater quality in El Eulma area, Algeria. *Appl Water Sci* 2:127–133
- Belkhiri L, Mouni L, Tiri A (2011) Water-rock interaction and geochemistry of groundwater from the Ain Azel aquifer, Algeria. *Environ Geochem Health*. doi:10.1007/s10653-011-9376-4
- Brown SD (1988) Chemometrics: a textbook. In: Massart DL, Vandeginste BGM, Deming SN, Michotte Y, Kaufman L (eds) Elsevier, Amsterdam, *J Chemom* 2:298–299, ISBN 0-444-42660-4. Price Dfl 175.00
- Chapman D (1992) Water quality assessment—a guide to use of biota, sediments and water in environmental monitoring (2nd Ed.). Chapman on Behalf of UNESCO, WHO and UNEP. Chapman & Hall, London
- Childs D (1970) The Essentials of Factor Analysis. Holt, Rinehart and Winston, [SI]
- Da Silva AMM, Sacomani LB (2001) Using chemical and physical parameters to define the quality of Pardo river water (Botucatu-SP-Brazil). *Water Res* 35:1609–1616
- Davis JC (1973) Statistics and data analysis in geology. Wiley, New York
- Dixon W, Chiswell B (1996) Review of aquatic monitoring program design. *Water Res* 30:1935–1948
- Evansa CD, Daviesa TD, Wigington PJ et al (1996) Use of factor analysis to investigate processes controlling the chemical composition of four streams in the Adirondack Mountains, New York. *J Hydrol* 185:297–316
- Gotz R, Steiner B, Friesel P, Roch K, Walkow F, Maa V, Reincke H, Stachel B (1998) Dioxin (PCDD/F) in the river elbe—investigations of their origin by multivariate statistical methods. *Chemosphere* 37:1987–2002
- Grande JA, González A, Beltrán R et al (1996) Application of factor analysis to the study of contamination in the aquifer system of Ayamonte-Huelva (Spain). *Groundw* 34:155–161
- Guler C, Thyne G, McCray J et al (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol J* 10:455–474
- Hooper RP, Peters NE (1989) Use of multivariate analysis for determining sources of solutes found in wet atmospheric deposition in the United States. *Environ Sci Technol* 23:1263–1268
- Karamouz M, Mahjouri N, Kerachian R (2004) River water quality zoning: a case study on Karoon and Dez river system. *Iran J Environ Eng Sci* 1(2):15–21
- Kuppusamy MR, Giridhar VV (2006) Factor analysis of water quality characteristics including trace metal speciation in the coastal environmental system of Chennai Ennore. *Environ Int* 32:174–179
- Liu CW, Lin KH, Kuo YM (2003) Application of factor analysis in the assessment of groundwater quality in a Blackfoot disease area in Taiwan. *Sci Total Environ* 313:7789
- Mendiguchía C, Moreno C, García-Vargas M (2007) Evaluation of natural and anthropogenic influences on the Guadalquivir River (Spain) by dissolved heavy metals and nutrients. *Chemosphere* 69:1509–1517





- Milanovic P (2004) Mission report, Watertightness of reservoir and dam site, Gotvand dam project, p 21
- Miller WR, Drever JI (1977) Water chemistry of a stream following a storm, Absaroka Mountains, Wyoming. *Geol Soc Am Bull* 88: 26–290
- Nakano T, Tayasu I, Yamada Y et al (2008) Effect of agriculture on water quality of Lake Biwa tributaries, Japan. *Sci Total Environ* 389:132–148
- Ouyang Y (2005) Evaluation of river water quality monitoring stations by principal component analysis. *Water Res* 39: 2621–2635
- Parinet B, Lhote A, Legube B (2004) Principal component analysis: an appropriate tool for water quality evaluation and management-application to a tropical lake system. *Ecol Model* 178: 295–311
- Pesce SF, Wunderlin DA (2000) Use of water quality indices to verify the impact of Cordoba City (Argentina) on Suquia River. *Water Res* 34:2915–2926
- Puckett LJ, Bricker OP (1992) Factors controlling the major ion chemistry of streams in the blue ridge and valley and ridge physiographic provinces of Virginia and Maryland. *Hydrol Process* 6:79–97
- Razmkhah H, Abrishamchi A, Torkian A (2010) Evaluation of spatial and temporal variation in water quality by pattern recognition techniques: a case study on Jajrood River (Tehran, Iran). *J Environ Manage* 91:852–860
- Reeder SW, Hitchon B, Levinson AA (1972) Hydro geochemistry of the surface waters of the Mackenzie River drainage basin, Canada. Factors controlling inorganic composition. *Geochim Cosmochim Acta* 36:825–865
- Reghunath R, Murthy TRS, Raghavan BR (2002) The utility of multivariate statistical techniques in hydro geochemical studies: an example from Karnataka, India. *Water Res* 36:2437–2442
- Shrestha S, Kazama F, Nakamura T (2008) Use of principal component analysis, factor analysis and discriminant analysis to evaluate spatial and temporal variations in water quality of the Mekong River. *J Hydroinformatics* 10:43–56
- Simeonov V, Stratis JA, Samara C et al (2003) Assessment of the surface water quality in Northern Greece. *Water Res* 37: 4119–4124
- Voutsas D, Zachariadis G, Samara C et al (1995) Evaluation of chemical parameters in Aliakmon river/Northern Greece. Part II: dissolved and particulate heavy metals. *J Environ Sci Health Part A Environ Sci Eng Toxicol* 30:1–13
- Wang YS, Lou ZP, Sun CC et al (2006) Multivariate statistical analysis of water quality and phytoplankton characteristics in Daya Bay, China, from 1999 to 2002. *Oceanologia* 48:193–213
- Wenning RJ, Erickson GA (1994) Interpretation and analysis of complex environmental data using chemometric methods. *TrAC Trends Anal Chem* 13:446–457
- Zarei H, Akhondali AM, Damough NA (2005) Study of impress Gachsaran Formation on Karoon River water quality in Khozestan province and compare with Dez River. 7th International River engineering conference. Shahid Chamran University of Ahvaz, Ahvaz, Iran (In persian)